基于改进退化隐马尔可夫模型的设备健康诊断与寿命预测研究 *

刘文溢，刘勤明†，叶春明，李冠林

(上海理工大学 管理学院, 上海 200093)

摘 要：针对隐马尔可夫模型在进行设备健康诊断时与实际存在较大偏差的问题，提出了一种以似幂关系加速退化为核心的改进退化隐马尔可夫模型(DGHMM)。首先，引入退化因子描述设备衰退过程，提出的似幂关系加速退化较常规指数式加速退化而言，能更好地描述设备服役期间随着役龄增加性能的逐步下降。其次，以全局搜索能力相对较强的改进遗传算法代替常规 EM 算法进行参数估计，克服了 EM 算法易陷入局部最优的局限性。同时，针对隐马尔可夫模型时间上须服从指数分布而不能直接用以寿命预测的局限性问题，提出了一种以近似算法与 Viterbi 算法为基础的贪婪近似法，以寻求最大概率剩余观测为目的，动态地寻求最大概率剩余状态路径，对设备剩余寿命进行预测。最后，通过美国卡特彼勒公司液压泵数据集对所提出的方法进行验证评价。结果表明，基于改进退化隐马尔科模型的设备健康诊断与寿命预测方法在描绘设备退化、设备状态诊断准确率方面更加有效，在剩余寿命预测上亦为可行。

关键词：隐马尔可夫模型；设备退化；健康诊断；剩余寿命预测；遗传算法；近似算法

中图分类号：TP306 doi: 10.19734/j.issn.1001-3695.2020.02.0067

Equipment health diagnostics and prognostics method based on improved degenerated HMM

Liu Wenyi, Liu Qinming, Ye Chunming, Li Guanlin

(*Business School*, *University of Shanghai for Science & Technology*, *Shanghai* 200093, *China*)

Abstract: In order to solve the problem of large deviation between Hidden Markov Model and actual equipment health diagnosis, this paper developed an improved Degenerated Hidden Markov Model (DGHMM) with a core of the quasi power relation. First, the model adopted the degradation factors, modeling the process of recession for the equipment's continuous decrease in performance. Compared with the conventional exponential accelerated degradation, the quasi power relation accelerated degradation can better describe the process that the performance of the equipment decreases gradually with the increase of service age. Then, the improved genetic algorithm can replace the conventional EM algorithm for parameters' estimation, which overcomes the limitation that the EM algorithm is easy to fall into local optimization. At the same time, in terms of the limitation of life prediction problem as a result of the Hidden Markov Model must obey exponential distribution, an algorithm named greed & approximation based on approximation algorithm and Viterbi algorithm came out, and to seek maximum probability remaining observation, for the purpose of seeking maximum probability dynamically surplus state path, to predict the residual life of equipment. Finally, the proposed method is validated and evaluated with the data set of caterpillar hydraulic pumps. The results show that the method of equipment health diagnosis and life prediction based on the improved degraded hidden Markov model is more effective in describing equipment's degeneration and the accuracy of equipment state diagnosis, and is also feasible in the prediction of residual life.

Key words: hidden Markov model; degradation health diagnostics; residual useful lifetime; genetic algorithm; approximation algorithm

0 引言

隐马尔可夫模型(hidden Markov model，HMM)作为一种统计学方法，从概率角度描述系统外在表现与隐含状态之间的关系，既能反映对象的随机性，又能反映其潜在结构，适用于处理复杂系统的建模问题，不仅在语音识别及分词方面取得了应用[1, 2]，在模式识别与健康诊断方面也取得了良好的效果。

Carey[3]等通过分析机械设备的振动信号首次将 HMM 成功应用于故障诊断领域；Wang[4]等建立了包含耦合的离散时变时滞和耦合的分布时变时滞的模型运用有效技术约束不确定项，并通过数值仿真验证了方法有效性。Ma[5]等阐述了 HMM 在时间上存在必须强制符合指数分布的不合理性而不能直接用以剩余寿命预测。总体而言，现阶段 HMM 模型的研究因其自身限制也仅停留在设备健康诊断方面。

为了克服 HMM 模型的局限性，学者们提出了隐半马尔可夫模型(hidden semi Markov model, HSMM)，Dong[6]等运用 HSMM 的状态驻留对设备进行了健康诊断与剩余寿命预测；Liu[7]运用 K-means 算法及交叉验证算法确定电网退化阶段，并用 HSMM 对电网进行诊断及寿命预测；除此之外对基于多传感器，多输入的模型，学者们也提出了一系列方法整合传感器信号，使得模型能够更准确地进行系统的状态识别及寿命预测[8, 9]。上述研究有着一个共同点，即 HSMM 本身依赖于先验分布，从而不可避免地存在一定程度的主观性特征。

随着对 HMM 及 HSMM 的研究逐渐深入，研究者们发

收稿日期：2020-02-03；**修回日期：**2020-03-23 **基金项目：**国家自然科学基金资助项目(71632008，71840003)；上海市自然科学基金资助项目(19ZR1435600)；教育部人文社会科学研究规划基金资助项目(20YJAZH068)

作者简介：刘文溢(1997-)，男，四川巴中人，硕士研究生，主要研究方向为故障诊断与寿命预测；刘勤明(1984-)，男(通信作者)，山东日照人，副教授，博士，主要研究方向为维护调度，人工智能等(lqm0531@163.com)；叶春明(1964-)，男，安徽宣城人，教授，博导，博士，主要研究方向为生产调度；李冠林(1998-)，男，山西大同人，本科生，主要研究方向为供应链管理.

现马尔可夫类模型对于状态的识别与实际依然存在较大的差异，为了得到更符合设备实际情况的模型，设备退化理论应运而生。Chen[10]等指出传统模型状态转移概率是固定的，但实际上随着设备工作时间的增加，设备损耗逐渐积累，设备的状态转移必然发生变化；Peng[11]设计了三类退化因子加入到 HSMM 模型当中，并对加入退化因子的模型的算法进行了重新推导，对参数进行了重估；He[12]等将三类退化因子运用到 HSMM 模型，并搭建柱塞泵实验台，通过实例验证了退化的有效性。上述对于设备退化的研究，大多是将退化应用到模型中并假设不同的先验分布，而对于退化因子本身的改进优化的研究还暂时空缺。

极大似然估计作为参数估计的常用方法，通过寻找使实验数据出现似然值最大的参数，来对一些复杂模型中的未知参数进行合理估计。相应地，EM 算法以其相对简单，效率高的特点而被广泛应用于参数估计，但 EM 算法本身却具有易陷入局部最优的缺点。为了寻求拓展，越来越多的研究者将智能算法应用到参数估计领域。Yuan[13]等人用改进 PSO 算法对 Baum-Welch 算法进行修正，获得了更加精确的参数；Yu[14]等将遗传算法运用到灰色模型中进行参数估计，数值结果显示，采用 GA 可大大提高模型预测精度；Krause[15]等采用基于遗传算法的函数逼近器估计感应电机的参数，并取得了较其他方法更好的效果；Zhang[16]等运用人工免疫算法对 HMM 模型进行优化，得到了辨识度最高的初始观测矩阵；Zhang[17]等运用自适应基因粒子群算法优化隐马尔可夫模型(CHMM)，得到了比传统 HSMM 更精确的结果。遗传算法作为全局搜索能力较强的智能寻优算法，尚未应用到马尔可夫类模型中，存在着可延展开拓的领域。

基于上述问题，首先引入退化因子，对设备衰退进行建模，得到设备退化过程的描述，并提出了一种以似幂关系加速退化为核心的 DGHMM，较常规指数式加速退化而言，能更加准确地描述设备性能随役龄增加而逐渐加速下降的过程。以改进遗传算法代替常规 EM 算法对整个模型进行参数估计；同时，针对 HMM 模型不能直接用以寿命预测的局限性问题，提出了一种基于近似算法与 Viterbi 算法的贪婪近似法，以设备退化时状态转移概率矩阵的变化为载体去预测设备剩余寿命。最后，以美国卡特彼勒公司液压泵数据集对模型进行验证评价，结果显示，本文方法是可行的和有效的。

1　HMM 模型

HMM 模型的表达式可描述为 $\lambda=(\pi,A,B)$，模型参数描述如下：

a) N：状态数，状态所构成的集合为 $\{S_1,S_2,S_3\cdots S_{N-1},S_N\}$，在任一时刻 t 所处的状态为 S_t，$S_t\in(s_1,s_2,s_3\cdots s_{N-1},s_N)$。

b) M：观测数，观测所构成的集合为 $\{\theta_1,\theta_2,\theta_3\cdots\theta_{M-1},\theta_M\}$，任一时刻 t 状态 i_t 产生的观测为 $o_t\in(\theta_1,\theta_2,\theta_3\cdots\theta_{M-1},\theta_M)$。

c) π：初始状态概率分布，$\pi=(\pi_1,\pi_2,\pi_3\cdots\pi_{N-1},\pi_N)$，其中所有概率之和为 1，初始概率向量的意义为 HMM 开始时，即 $t=1$ 时，模型处于第 $n(n\in(1,2,3\cdots N-1,N))$ 个状态的概率。

d) A：状态转移矩阵。A 是一个 $N\times N$ 的矩阵，$A=[\alpha_{ij}]_{N\times N}$，$\alpha_{ij}$ 为从状态 i 转移到状态 j 的概率，即 $\alpha_{ij}=P(S_{t+1}=j\mid S_t=i)$ 且 $\sum_{j=1}^{N}\alpha_{ij}=1(i\in(1,2,3\cdots N-1,N))$，$\sum_{i=1}^{N}\alpha_{ij}=1(j\in(1,2,3\cdots N-1,N))$。

e) B：观测概率分布。$B=[b_j(K)]_{N\times M}$，其中 $b_j(K)=P(o_t=\theta_K\mid S_t=j)(j\in(1,2,3\cdots N),K\in(1,2,3\cdots M))$，且 $\sum_{k=1}^{M}b(k)=1$。

2　设备退化过程

2.1　退化因子设计

传统 HMM 假定在设备整个寿命周期内，其状态转移发生的概率是不变的，而这个假设限制了 HMM 实际建模分析中的能力。对此，国内外学者提出了时变 HMM。典型的设备退化规律可由图 1 所示。整个退化过程可分为 A、B、C 三个阶段，其中 A 段为平稳退化阶段，此阶段设备性能指数基本呈现稳定态势；B 段为均匀退化阶段，设备性能指数呈现均匀变化；C 段为加速阶段，设备性能随着寿命的推移呈现加速下降的趋势，健康状况急剧恶化，直至设备性能指标达到完全失效阈值。

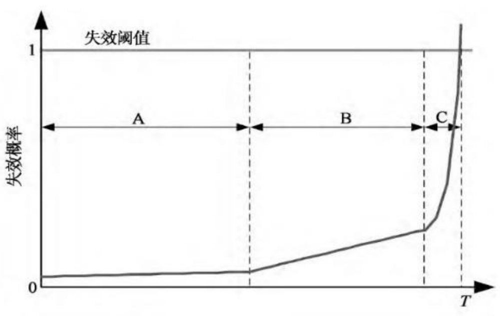


图 1　典型设备退化曲线

Fig. 1　Typical equipment degradation curve

在现有大多数研究中，国内外学者们普遍将 A、B、C 三阶段退化过程的退化因子分别定义为常数，线性函数以及指数函数。对于指数退化阶段，普遍将相邻两时段 t 与 $t+\Delta t$ 的状态内转移概率表示为

$$\alpha_{ii}(t+1)=\alpha_{ii}(t)^{\phi}\quad i\in(1,2,3\cdots N)$$

其中 ϕ 为待估计参数，为定值。基于 Peng[11]对退化因子的定义，整个设备加速退化为一递归过程，对于离散时间的状态转移，定义状态退化关系式为 f，递归 X 次后的内转移概率值为 f^{x}，则退化函数应满足：

$$f^{0}(i)=a_{ii}(t_{end})\tag{1}$$

$$f^{X}-f^{X-1}>f^{X-1}-f^{X-2}\quad(X\ge 2)\tag{2}$$

式(1)表示刚从均匀退化状态结束时设备状态内转移概率，即递归的初值，式(2)表示加速退化过程中加速的概念，即越靠近最终寿命时，设备停留在当前状态的概率下降越快。

针对加速退化关系式的定义，在分析指数递归关系式结构的基础上，提出了一种新的似幂关系式加速退化关系：

$$\alpha_{ii}(t+1)=\phi^{\alpha_{ii}(t)}-1\qquad i\in(1,2,3\cdots N)\tag{3}$$

其中 ϕ 为待估计参数，可由历史数据估计得出，由于状态转移概率 $\alpha_{ii}(t)$ 恒小于 1，ϕ^{0} 为 1，则需要减去边界值得到适应后的状态转移概率。

对于选取满足递归加速退化要求的相适应参数进行指数关系与似幂关系比较，发现合理参数下的似幂关系与指数关系有着相似的递归加速退化效果，如图 2 所示。

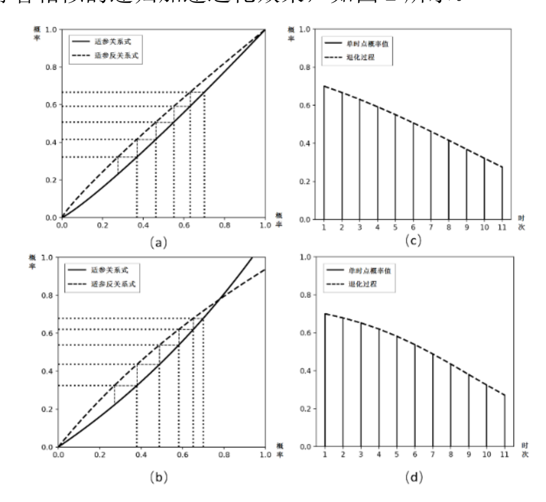


图 2　关系式加速退化递归示意图

Fig. 2　Schematic diagram of relational accelerated degradation recursion

其中子图(a)(b)分别为满足加速退化过程且在合理参数下的指数关系及似幂关系退化示意图，粗实线为关系式的连续函数图，粗虚线为对应关系式的反函数图，点线为递归过程中状态转移概率，由图可知两种关系式具有一定程度的作用相似性。子图(c)(d)分别为当前示例合理参数下的指数关系退化过程以及似幂关系退化过程，实线表示各个单时点状态概率，虚线表示当前关系式下概率的部分退化过程。由图看出，似幂关系在表达加速退化过程中，程度优于指数关系。

2.2　退化转移概率推算

a) 平稳退化阶段，相邻两时点 t 与 $t+\Delta t$ 设备保持当前状态概率的变化为定值，即

$$\alpha_{ii}(t+1)-\alpha_{ii}(t)=\phi_1 \quad i\in(1,2,3\cdots N)$$

其中 $\phi_1\le 0$，保持当前状态概率 $\alpha_{ii}(t)$ 减小，则由当前状态转移至其他状态的概率会随之增大，且根据比例分配给各 $\alpha_{ij}(t)\ (j=i+1,i+2\cdots N)$。则平稳退化状态下一时刻状态转移概率为

$$\begin{cases}\alpha_{ii}(t+1)=\alpha_{ii}(t)+\phi_1\\ \alpha_{ij}(t+1)=\alpha_{ij}(t)-\dfrac{\alpha_{ij}(t)\phi_1}{\sum_{j=i+1}^{N}\alpha_{ij}(t)}\end{cases} \tag{4}$$

则整个平稳退化过程某一时点的状态转移概率可由初始时刻递推得到：

$$\begin{cases}\alpha_{ii}(k\Delta t)=\alpha_{ii}(0)+k\phi_1\\ \alpha_{ij}(t+1)=\alpha_{ij}(t)+\dfrac{\alpha_{ij}(0)k\phi_1}{\sum_{j=i+1}^{N}\alpha_{ij}(0)}\end{cases} \tag{5}$$

b) 均匀退化阶段，相邻两时段 t 与 $t+\Delta t$ 设备保持当前状态的概率的变化为 t 时段概率的线性值，即

$$\alpha_{ii}(t+1)-\alpha_{ii}(t)=\phi_2\ \alpha_{ii}(t) \quad i\in(1,2,3\cdots N)$$

其中 $\phi_2\le 0$，$\alpha_{ii}(t)$ 减少的值同样分配给 $\alpha_{ij}(t)$：

$$\begin{cases}\alpha_{ii}(t+1)=(\phi_2+1)\alpha_{ii}(t)\\ \alpha_{ij}(t+1)=\alpha_{ij}(t)+\dfrac{\alpha_{ij}(t)\phi_2\alpha_{ii}(t)}{\sum_{j=i+1}^{N}\alpha_{ij}(t)}\end{cases} \tag{6}$$

若定义平稳退化阶段结束时的时点为 t_{e1}，则处于均匀退化时的某一时点的状态转移概率可描述为

$$\begin{cases}\alpha_{ii}(t_{e1}+k\Delta t)=(\phi_2+1)^k\alpha_{ii}(t_{e1})\\ \alpha_{ij}(t_{e1}+k\Delta t)=\alpha_{ij}(t_{e1})+\dfrac{\alpha_{ij}(t_{e1})((\phi_2+1)^k-1)\alpha_{ii}(t_{e1})}{\sum_{j=i+1}^{N}\alpha_{ij}(t_{e1})}\end{cases} \tag{7}$$

c) 对于本文提出的基于幂函数的加速退化阶段，相邻两时点 t 与 $t+\Delta t$ 设备保持当前状态的概率的变化为 t 的似幂关系

$$\alpha_{ii}(t+1)=\phi_3^{\alpha_{ii}(t)}-1 \qquad i\in(1,2,3\cdots N)$$

其中 $\phi_2>0$，$\alpha_{ii}(t)$ 减少的值分配给 $\alpha_{ij}(t)$，则下一时刻的状态转移概率为

$$\begin{cases}\alpha_{ii}(t+1)=\phi_2^{\alpha_{ii}(t)}-1\\ \alpha_{ij}(t+1)=\alpha_{ij}(t)+\dfrac{\alpha_{ij}(t)(\phi_2^{\alpha_{ii}(t)}-1-\alpha_{ii}(t))}{\sum_{j=i+1}^{N}\alpha_{ij}(t)}\end{cases} \tag{8}$$

同样定义均匀退化阶段结束时点为 t_{e2} 由于似幂关系特殊的递推形式，引用递归关系符号 $f^x(x)$，其中 X 表示递归次数，x 表示递归概率初值，则在此加速退化阶段某一时点状态转移概率为

$$\begin{cases}\alpha_{ii}(t+k\Delta t)=f^k(\alpha_{ii}(t_{e2}))\\ \alpha_{ij}(t+k\Delta t)=\alpha_{ij}(t_{e2})+\dfrac{\alpha_{ij}(t_{e2})(f^k(\alpha_{ii}(t_{e2}))-\alpha_{ii}(t_{e2}))}{\sum_{j=i+1}^{N}\alpha_{ij}(t_{e2})}\end{cases} \tag{9}$$

通过设备历史数据训练后得到初始状态转移概率矩阵 A，因设备的运转退化过程在无人工干预的自然状态下是不可逆的过程，即设备状态只由停留在当前状态及由当前状态退化到更差状态，因此初始状态转移概率矩阵 A 可描述为

$$A=\begin{vmatrix}\alpha_{11}&\alpha_{12}&\cdots&\alpha_{1N-1}&\alpha_{1N}\\ 0&\alpha_{22}&\cdots&\alpha_{2N-1}&\alpha_{2N}\\ \vdots&\vdots&\ddots&\vdots&\vdots\\ 0&0&\cdots&\alpha_{N-1N-1}&\alpha_{N-1N}\\ 0&0&\cdots&0&1\end{vmatrix}$$

结合上述三种退化递推式及转移概率矩阵则可推算出三个不同退化阶段中经历 $k\Delta t$ 后的状态转移概率矩阵为

$$A_{k\Delta t}=\begin{vmatrix}\alpha_{11}(0)+k\phi_1 & \alpha_{12}(0)+\dfrac{\alpha_{12}(0)k\phi_1}{\sum_{j=1+1}^{N}\alpha_{1j}(0)} & \cdots & \alpha_{1N}(0)+\dfrac{\alpha_{1N}(0)k\phi_1}{\sum_{j=1+1}^{N}\alpha_{1j}(0)}\\ 0 & \alpha_{22}(0)+k\phi_1 & \cdots & \alpha_{2N}(0)+\dfrac{\alpha_{1N}(0)k\phi_1}{\sum_{j=1+1}^{N}\alpha_{1j}(0)}\\ 0&0&\ddots&\vdots\\ 0&0&\cdots&1\end{vmatrix} \tag{10}$$

$$(0\le k\le \frac{t_{e1}}{\Delta x})$$

$$A_{k\Delta t}=\begin{vmatrix}(\phi_2+1)^{k-\frac{t_{e1}}{\Delta x}}\alpha_{11}^{e1} & \alpha_{12}^{e1}+\dfrac{\alpha_{12}^{e1}((\phi_2+1)^{k-\frac{t_{e1}}{\Delta x}}-1)\alpha_{11}^{e1}}{\sum_{j=2}^{N}\alpha_{1j}^{e1}} & \cdots & \alpha_{1N}^{e1}+\dfrac{\alpha_{1N}^{e1}((\phi_2+1)^{k-\frac{t_{e1}}{\Delta x}}-1)\alpha_{11}^{e1}}{\sum_{j=2}^{N}\alpha_{1j}^{e1}}\\ 0 & (\phi_2+1)^{k-\frac{t_{e1}}{\Delta x}}\alpha_{22}^{e1} & \cdots & \alpha_{2N}^{e1}+\dfrac{\alpha_{2N}^{e1}((\phi_2+1)^{k-\frac{t_{e1}}{\Delta x}}-1)\alpha_{22}^{e1}}{\sum_{j=2}^{N}\alpha_{2j}^{e1}}\\ 0&0&\ddots&\vdots\\ 0&0&\cdots&1\end{vmatrix} \tag{11}$$

$$(\frac{t_{e1}}{\Delta x}\le k\le \frac{t_{e2}}{\Delta x})$$

$$A_{k\Delta t}=\begin{vmatrix}f^{k-\frac{t_{e1}}{\Delta x}}(\alpha_{11}^{e2}) & \alpha_{12}^{e2}+\dfrac{\alpha_{12}^{e2}(f^k(\alpha_{12}^{e2})-\alpha_{12}^{e2})}{\sum_{j=2}^{N}\alpha_{1j}^{e2}} & \cdots & \alpha_{1N}^{e2}+\dfrac{\alpha_{1N}^{e2}(f^k(\alpha_{1N}^{e2})-\alpha_{1N}^{e2})}{\sum_{j=2}^{N}\alpha_{1j}^{e2}}\\ 0 & f^{k-\frac{t_{e1}}{\Delta x}}(\alpha_{22}^{e2}) & \cdots & \alpha_{2N}^{e2}+\dfrac{\alpha_{2N}^{e2}(f^k(\alpha_{2N}^{e2})-\alpha_{2N}^{e2})}{\sum_{j=3}^{N}\alpha_{2j}^{e2}}\\ 0&0&\ddots&\vdots\\ 0&0&\cdots&1\end{vmatrix} \tag{12}$$

$$(\frac{t_{e2}}{\Delta x}<k)$$

上述式(10)表示从初始时点至 t_{e1} 时点设备系统处于平稳退化阶段，经过 $k\Delta t$ $(0\le k\le \frac{t_{e1}}{\Delta x})$ 时的状态转移概率矩阵，描述为 $A_{k\Delta t}$。将当 $k=\frac{t_{e1}}{\Delta x}$ 时的状态转移概率矩阵记作 A_{e1}，描述为 $A_{e1}=(\alpha_{ij}^{e1})_{N\times N}$，其中 α_{ij}^{e1} 为平稳退化阶段结束时状态转移概率矩阵的第 i 行，第 j 列的元素。则根据描述，可得式(11)所示设备处于均匀退化状态，且经历 $k\Delta t$ $(\frac{t_{e1}}{\Delta x}<k\le\frac{t_{e2}}{\Delta x})$ 时的状态转移概率矩阵。同样的，将当 $k=\frac{t_{e2}}{\Delta x}$ 时的状态转移概率矩阵记作 A_{e2}，描述为 $A_{e1}=(\alpha_{ij}^{e2})_{N\times N}$，其中 α_{ij}^{e2} 为均匀退化阶段结束时状态转移概率矩阵第 i 行，第 j 列的元素，由此得式(12)所示系统设备处于似幂关系加速退化状态并经历了 $k\Delta t$ $(\frac{t_{e2}}{\Delta x}<k)$ 的状态转移概率矩阵。

3　基于改进遗传算法的参数估计

对于极大似然估计常用的 EM 算法容易陷入局部最优的缺陷，本文采用了全局搜索能力更强的遗传算法进行 DGHMM 的参数估计，建立对数似然函数

$$\log(L(\lambda,\phi_1,\phi_2,\phi_3))=\log(\Pr(\theta_1,\theta_2\cdots\theta_T\mid\lambda,\phi_1,\phi_2,\phi_3))$$

将每组用以迭代的参数组计算而出的似然值作为单个解决方案的适应度，则每一次种群迭代的最优个体便可表示为

$$\overline{\lambda},\overline{\phi_1},\overline{\phi_2},\overline{\phi_3}=\arg\max_{\lambda,\phi_1,\phi_2,\phi_3}\log(L(\lambda,\phi_1,\phi_2,\phi_3))$$

$$=\arg\max_{\lambda,\phi_1,\phi_2,\phi_3}\log(\Pr(\theta_1,\theta_2\cdots\theta_T\mid\lambda,\phi_1,\phi_2,\phi_3))$$

在经过若干次迭代后发现 $L(\lambda,\phi_1,\phi_2,\phi_3)$ 值变化不大或达到预定迭代次数后，便输出目前较优的参数组 $\lambda,\phi_1,\phi_2,\phi_3$。

3.1　改进编码

本文采用混合编码策略：对于转移概率矩阵 $A(N\times N)$ 以及观测矩阵 $B(N\times M)$，首先对矩阵进行打平，由于转移概率矩阵自身的特性，则其矩阵中实际有价值的数值仅有 $\frac{N^2+N-2}{2}$ 个。考虑到概率值均小于 1，且大部分为多位数的小数，采用十进制浮点数编码，且保留三位小数。打平后的转移概率矩阵去零去一后便是一个参数组中转移概率染色体，打平后的观测概率矩阵便是参数组中观测染色体。初始状态分布 π 同样使用 M 个位点的十进制浮点数编码染色体表示(保留三位小数)。

对于 ϕ_1,ϕ_2,ϕ_3，考虑到整个模型对退化因子极为敏感，采用二进制的整数编码且保留五位小数。以伪 18 个位点的二进制整数染色体对 ϕ_1,ϕ_2,ϕ_3 的绝对值进行转换，其中第一个位点表示参数整数部分，且该位点不参与二进制与十进制的相互转换。相应地，余下的十七位在解码过程中需要缩小十万倍，然后加上首位编码数值，便得到解码后的参数。各变量编码后的染色体如图 3 所示。

λ 三参的编码中，染色体一个位点代表了矩阵中的一个概率值，每个位点由四个部分组成。灰色部分为矩阵中单个值的整数部分，白色部分分别为该值的十分位、百分位以及千分位的值。在整个改进遗传算法流程中，单个参数组里的参数总是同时参与计算，一个参数组即为一个完整的解决方案。

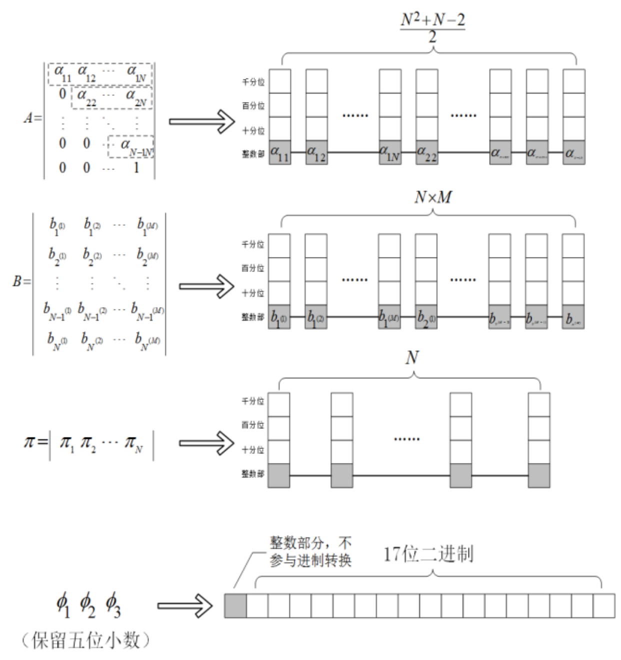


图 3　编码示意图

Fig. 3　Coding diagram

3.2　改进交叉

对于上述改进过后的混合编码策略，相应地对交叉与变异方式进行改进。针对编码方式相似的 λ 三参数，设计两种不同的交叉方式 1 与交叉方式 2。交叉方式 1 是交换随机两个染色体相同位点上千、百分位数和十分、整数位数，并对产生的两条新的染色体各自进行自适应归一化；交叉方式 2 是交换随机两条染色体各自一部分位点，并对产生的新染色体进行自适应归一化，交叉方式 1、2 分别如图 4、5 所示。

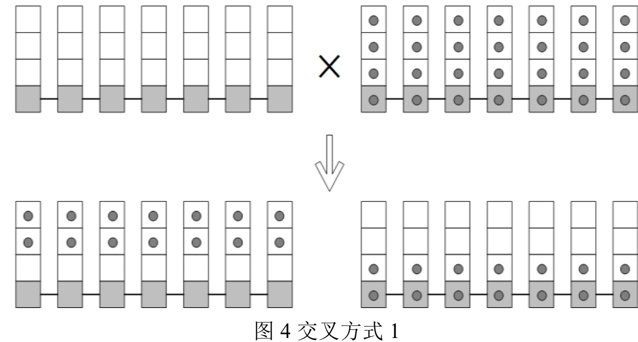


图 4 交叉方式 1

Fig. 4　Crossover mode 1

图 5 交叉方式 2

Fig. 5　Crossover mode 2

由于染色体是概率矩阵打平后的产物，矩阵同行元素和为 1 决定了自适应归一化需要满足一定范围限制，具体的自适应归一化规则如表 1 所示，染色体位点从 1 开始计数。

表 1　自适应归一规则

Tab. 1　Adaptive normalization rule

染色体类别	归一化	
	组别	归一范围
转移染色体(N*N)	1	1—N
	2	N+1—2N-1
	……	……
	N-1	$\frac{N^2+N-4}{2}-\frac{N^2+N-2}{2}$
观测染色体(N*M)	1	1—M
	2	M+1—2M
	……	……
	N	(N-1)M+1—NM
初始染色体(N)	1	1—N

λ 三参的变异以及退化因子的交叉变异均同常规 GA 算法，变异操作后的归一化同样遵从表 1 规则。

改进后的交叉操作，利用两种不同的交叉方式，通过随机选择交叉方式，极大程度地增加了种群的多样性，增强了算法的全局搜索能力，同时在一定程度上保留了算子精度，有利于产生全局最优参数。

在算法完整运行后，使得似然函数取值最大的参数组 $\lambda,\phi_1,\phi_2,\phi_3$ 便是历史数据条件下，HMM 模型最优的估计参数。

4　基于 DGHMM 的设备剩余寿命预测

4.1　贪婪近似法

对于常规 HMM 模型的预测问题，一般存在近似算法以及 Viterbi 算法。现阶段研究中，普遍在得到 HMM 模型参数以及观测序列后运用后者计算此观测序列下概率最大的状态序列，即

$$\bar{I}=\arg\max\Pr(\theta_1,\theta_2\cdots\theta_T\mid\lambda,I)$$

实际应用中，在需要预测设备系统剩余寿命时，往往设

备还处于服役中，即此时是得不到设备从服役初期开始至失效的整个过程的观测值。为了解决这一问题，研究者们一般用同种设备历史数据加以代替并建立模型，输入服役设备的某一时点之前的观测值序列，预测其隐状态，再通过役龄时间分布来确定剩余寿命，但这与 HMM 模型必须服从指数分布的局限性冲突；且设备之间存在的固有差异以及设备状态的随机性，使得预测在一定程度上损失设备适应性。

因此本文在近似算法以及 Viterbi 算法基础上提出一种预测剩余状态的新方法(下称贪婪近似法)。

贪婪近似法将近似算法寻求状态的思想，运用到寻找剩余观测上，即在已知 HMM 参数 λ 的前提下，寻求下一时刻最有可能出现的观测值。

定义设备在第 t 时刻产生观测 θ_k 的前提下，第 $t+1$ 时刻产生观测 θ_g 的概率为 $v_{kg}(t)=P(o_{t+1}=\theta_g\mid o_t=\theta_k,\lambda)$ $(k,g=1,2\cdots M)$，则 v 的数学描述为

$$v_{kg}(t)=\sum_{j=1}^{N}\left(b_j(g)\sum_{i=1}^{N}\left(\frac{b_i(k)\alpha_{ij}^{t}}{\sum_{i=1}^{N}b_i(k)}\right)\right)\quad(k,g=1,2\cdots M)\tag{13}$$

其中 α_{ij}^{t} 为 t 时刻设备系统的状态转移概率，其如上述所说符合三阶段退化代入不同的 k,g 值便能计算出从 t 时刻的 k 观测转移到 $t+1$ 时刻的 g 观测的概率，由此得出间接观测转移概率矩阵为

$$V^{t}=\begin{vmatrix}V_{11}^{t} & V_{12}^{t} & \cdots & V_{1M-1}^{t} & V_{1M}^{t}\\ V_{21}^{t} & V_{22}^{t} & \cdots & V_{2M-1}^{t} & V_{2M}^{t}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ V_{M-11}^{t} & V_{M-12}^{t} & \cdots & V_{M-1M-1}^{t} & V_{M-1M}^{t}\\ V_{M1}^{t} & V_{M2}^{t} & \cdots & V_{MM-1}^{t} & V_{MM}^{t}\end{vmatrix}$$

由于状态转移矩阵 A 随着监测时点逐渐退化，则 V 也同样是随着时间变化的。对于某一产生观测 $\theta_i(i=1,2\cdots M)$ 的时点 t，其下一时点 $t+1$ 最有可能出现的观测值概率为

$$P(t+1)=\max_{1\le j\le M}v_{ij}^{t}\,(i=1,2\cdots M)$$

下一时点最优可能观测则为

$$o_{t+1}(j)=\arg\max_{1\le j\le M}v_{ij}^{t}\,(i=1,2\cdots M)$$

则此时的观测序列为前 t 个实际观测值 $O_{old}=(o_1,o_2\cdots o_t)$ 和一个预测观测值 o_{t+1}^{*} 组成，即 $O_{new}=(o_1,o_2\cdots o_t,o_{t+1}^{*})$。

详细的剩余状态预测步骤如下：

a) 将 O_{old} 代入 Viterbi 算法求得最大概率状态路径 $Road_{old}$，并记录最后一个状态值为 s。

b) 运用近似法求得下一时点最有可能出现的观测 θ^{*}，并将其拼接到 O_{old} 之后形成 O_{new}。

c) 将 O_{new} 代入 Viterbi 算法求得新的最大概率状态路径 $Road_{new}$，并记录最后一个状态值为 s^{*}。

d) 判断 s 与 s^{*}，如果 $s=s^{*}$，则重复步骤 b)~d)；如果 $s\ne s^{*}$ 则至步骤 e)。

e) 最初的 O_{old} 与最终的 O_{new} 相减便是预测的剩余观测，输出剩余观测。

在求得剩余观测之后，运用标准 Viterbi 算法，代入加入剩余观测后的观测序列及模型参数，寻求当前子状态下概率最大剩余状态路径 $I_{\max}$。

4.2 设备剩余寿命预测

对于 4.1 节求得的最大剩余状态路径 $I_{\max}$，假设当前设备处于状态 i，则在 $I_{\max}$ 中找到处于状态 i 的最后一个时间节点 t_{last}，此时的状态转移概率矩阵中第 i 行的数据可由上述退化因子及初始状态转移概率矩阵推算而知，则设备的剩余寿命可表示为

$$\mathrm{RUL}=(t_{last}-t)\Delta t+\frac{a_{ii}^{last}\Delta t}{\sum_{j=i+1}^{N}a_{ij}^{last}}+\sum_{j=i+1}^{N}D_j\tag{14}$$

式(14)中 D_j 为历史数据中设备处于 j 状态的寿命，a_{ii}^{last} 为当前状态 i 最后一个节点时的由 i 状态转向 i 状态的概率值，a_{ij}^{last} 为当前状态 i 最后一个节点时的由 i 状态转向 j 状态的概率值。即设备总剩余寿命等于当前状态寿命余量以及后续所有状态的寿命之和。

5 算例分析

通过美国卡特彼勒公司液压泵的设备健康诊断与寿命预测实例来验证评价本文提出的模型与方法。实验室中设备的振动信号是由安装在与液压泵旋转轴平行位置的液压加速计收集。在应用实例中，分别对液压泵充入 20、40、60 与 80mg 的微尘，并每隔 10min 采集一个时间大约为 1min 的振动信号样本。随后使用 10db 的小波将振动信号分为五层，将经过降维后的小波系数作为 DGHMM 的输入特征序列向量。整个实验过程中，液压泵的状态可分为四种，分别为 Baseline、Cont1、Cont2 以及 Cont3，其中 Cont3 状态为设备的彻底失效状态。

整个实验分析平台为 Python3，平台运行环境为 Windows 10。

5.1 DGHMM 数据准备

利用来自 36 个传感器(Sensor)的振动信号监测数据，对液压泵进行健康状态诊断以及剩余寿命预测。部分经过小波变换后的振动数据(Pump6)如表 2 所示。

表 2 Pump6 部分小波变换数据

Tab. 2 Partial wavelet transform data of pump6

时点	Sen 1	Sen 7	Sen 15	Sen 23	Sen 32
1	2.62	19.53	5.66	0.06	0.96
3	2.65	20.24	6.10	0.06	1.04
5	2.39	17.45	4.92	0.06	0.87
7	2.07	16.15	3.98	0.06	0.75
9	2.25	18.53	4.51	0.06	0.81
11	2.24	21.44	4.38	0.06	0.87
13	2.41	5.37	6.84	0.08	1.66
15	2.58	6.16	7.96	0.09	1.84
17	2.53	6.07	7.63	0.09	1.84
19	2.44	6.24	7.71	0.10	1.84
21	20.66	9.68	8.42	0.11	2.17
23	5.06	8.75	7.20	0.10	1.89
25	7.53	8.82	8.15	0.10	2.00
27	5.74	8.37	7.19	0.10	1.89
29	3.46	8.25	7.16	0.10	1.92
31	41.78	7.98	7.12	0.15	1.86
33	31.08	7.90	6.98	0.14	1.84
35	60.15	7.99	7.03	0.15	1.85
37	77.27	7.80	7.12	0.17	1.84

5.2 实验参数估计

本文运用改进遗传算法同时对 DGHMM 与时变 HMM 进行参数估计。表 3 与表 4 分别为泵 6 估计参数组中初始与失效临界状态转移概率矩阵。算法中种群数量为 30，交叉概率为 0.7，变异概率为 0.3。考虑到不同参数组下实验数据出现的概率似然值可能存在数量级差异较大的情况，通过减少种群数量、增加迭代次数来增强选择力度，相应地通过增加变

异概率来保证每代种群多样性。

DGHMM 与时变 HMM 在改进 GA 算法整个流程中的概率似然值迭代曲线如图 6 所示(Pump6)。从参数估计的结果来看，实验数据在 DGHMM 的较优参数组下出现的似然值大于在时变 HMM 的较优参数组下出现的似然值。这一结果在一定程度上可反映 DGHMM 较常规基于指数式加速退化的时变 HMM 而言，能更好地描绘实验数据。

表 3　Pump6 初始状态转移矩阵

Tab. 3　Initial state transition matrix of pump6

state	baseline	cont1	cont2	cont3
baseline	0.924	0.064	0.007	0.005
cont1	0.000	0.783	0.128	0.089
cont2	0.000	0.000	0.807	0.193
cont3	0.000	0.000	0.000	1.000

表 4　Pump6 失效临界状态转移矩阵

Tab. 4　Failure critical state transition matrix of pump6

state	baseline	cont1	cont2	cont3
baseline	0.426	0.483	0.053	0.038
cont1	0.000	0.361	0.377	0.262
cont2	0.000	0.000	0.403	0.597
cont3	0.000	0.000	0.000	1.000

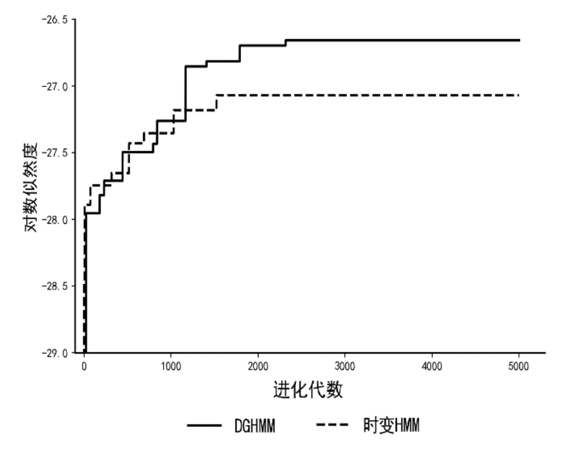


图 6　Pump6 似然迭代对比图

Fig. 6　Likelihood iteration comparison diagram of pump6

5.3　健康状态诊断

分别将泵 6 与泵 82 的实验数据向量化后输入 HMM、时变 HMM 以及 DGHMM 模型，运用 Viterbi 算法分别求得其在各自实验观测数据和各自较优参数组下设备的最大概率状态路径并加以验证评价，结果如表 5 所示，泵 6 与泵 82 实验数据在不同模型下的最大概率状态路径概率值如表 6 所示。

表 5　泵 6 与泵 82 健康状态识别结果

Tab. 5　Health states identification results of pump6 and pump82

类别	pump6						pump82					
	状态	base	cont1	cont2	cont3	准确率	状态	base	cont1	cont2	cont3	准确率
HMM	base	12	0	0	0	100%	base	9	0	0	0	100%
	cont1	1	6	0	0	86%	cont1	2	8	0	0	80%
	cont2	0	3	8	0	73%	cont2	0	4	6	0	60%
	cont3	0	0	0	8	100%	cont3	0	0	1	11	92%
时变 HMM	base	12	0	0	0	100%	base	9	0	0	0	100%
	cont1	0	7	0	0	100%	cont1	0	10	0	0	100%
	cont2	0	0	7	4	64%	cont2	0	0	7	3	70%
	cont3	0	0	0	8	100%	cont3	0	0	0	12	100%
DGHMM	base	12	0	0	0	100%	base	9	0	0	0	100%
	cont1	0	7	0	0	100%	cont1	0	10	0	0	100%
	cont2	0	0	11	0	100%	cont2	0	0	10	0	100%
	cont3	0	0	0	8	100%	cont3	0	0	0	12	100%

表 6　泵 6 与泵 82 在各模型下路径概率值

Tab. 6　The path probabilities of pump 6 and pump 82 under each model

类别	Pump6	Pump82
HMM	7.32E-11	8.16E-11
时变 HMM	5.39E-13	1.03E-13
DGHMM	1.36E-13	1.77E-12

5.4　剩余寿命预测

运用提出的贪婪近似算法以及常规 Viterbi 算法的设备剩余寿命预测方法，动态地求得 Pump6 与 Pump82 在 DGHMM 下的剩余寿命，结果由图 7 所示。从图中可以看出预测而出的剩余寿命值与实际剩余寿命的误差处于相对合理范围内，且当设备处于失效状态 cont3 时，设备停止运作，退化相应停止，模型也相应地成功识别并判定设备剩余寿命为 0，虽然预测值一定程度上在 Baseline 与 cont1 状态转换处存在“滞留现象”，但整体而言，基于 DGHMM 的贪婪近似法对于 Pump6 与 Pump82 剩余寿命的预测是有效可行的。

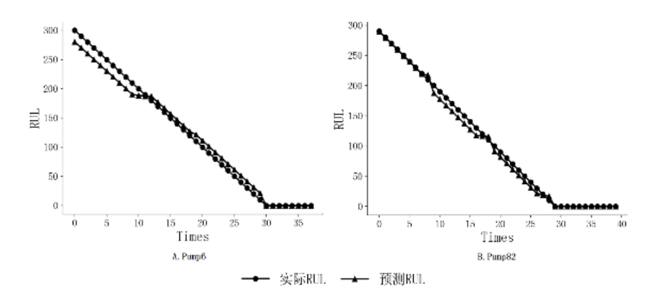


图 7　Pump6 与 pump82 剩余寿命预测图

Fig. 7　RUL of pump6 and pump82

6　结束语

本文提出了一种以新加速退化为核心的设备健康诊断与剩余寿命预测方法。提出的似幂关系更保守，较同为退化因子的指数式关系而言更稳定，且能更好描述设备性能随役龄增加的逐步加速退化的过程，对于设备的状态识别也更精确；基于改进遗传算法的参数估计方法，也取得了相对较好的效果；提出的贪婪近似法一定程度上打破了常规隐马尔可夫模型时间上必须服从指数分布的束缚，能够近似地对设备的剩余寿命进行预测。因此基于改进退化隐马尔可夫模型的设备健康诊断与剩余寿命预测方法在描绘设备退化、设备状态诊断准确率方面更加有效，在剩余寿命预测上亦为可行。

本文仅将似幂关系应用到了 HMM 模型并验证了结果，未来应将其应用到 HMM 的拓展——HSMM 中，进一步对似幂关系进行验证。为保证寻优效果，文中的改进遗传算法迭代次数较多，时间复杂度较高，未来也应尝试改善；且在贪婪近似法进行过程中若出现无限循环情况，还应设置相应的迭代次数限制才能完成算法相应步骤。

参考文献：

[1] 刘群，张华平，俞鸿魁，等. 基于层叠隐马模型的汉语词法分析 [J]. 计算机研究与发展, 2004, (08): 1421-1429. (Liu Qun, Zhang Huaping, Yu Hongkui, *et al.* Chinese Lexical Analysis Using Cascaded Hidden Markov Model [J]. Journal of Computer Research and Development, 2004, (08): 517-522.)

[2] 曹林，王东峰，刘小军，等. 基于 Gabor 小波和 HMM 的人像鉴别算法 [J]. 控制与决策, 2005, (09): 1073-1076. (Cao Lin, Wang Dongfeng, Liu Xiaojun, *et al.* Face Recognition Based on Gabor Wavelets and Hidden Markov Model [J]. Control and Decision, 2005, (09): 1073-1076.)

[3] Carey B, Dan M, Tarik A. Condition-based maintenance of machines using hidden markov models [J]. Mechanic Systems and Signal

Processing, 2000, 14 (4): 597-612

[4] 王燕锋，李祖欣，全立地，等．具有不确定转移概率的马尔可夫复杂网络的聚类同步 [J]．控制与决策，2018, 33 (04): 741-748. (Wang Yanfeng, Li Zuxin, Quan Lidi, *et al.* Cluster synchronization of Markovian complex networks with uncertain transition probabilities [J]. Control and Decision, 2018, 33 (04): 741-748)

[5] 马伦，康建设，赵强．基于 HMM 的设备剩余寿命预测框架及其实现 [J]．计算机仿真，2010, 27 (05): 88-91. (Ma Lun, Kang Jianshe, Zhao Qiang. Implementation of Equipment Residual Life Prediction Framework Based on Hidden Markov Model [J]. Computer Simulation, 2010, 27 (05): 88-91.)

[6] 董明．一种基于自回归隐式半 Markov 链的设备健康管理新方法 [J]．中国科学：信息科学，2008, (12): 2185-2198. (Dong Ming, A new device health management method based on autoregressive implicit semi-Markov chain [J]. Science in China: Information Sciences, 2008, (12): 2185-2198.)

[7] 刘浅．基于动静态参数检测的高速铁路接触网剩余寿命估计 [D]．四川：西南交通大学，2017. (Liu Qian. Residual Life Prediction Based on Dynamic and Static Parameters of High Speed-Railway Catenary [D]. Sichuan: Southwest Jiaotong University, 2017.)

[8] Dong M. Hidden semi-Markov model based method-ology for multi-sensor equipment health diagnosis and prognosis [J]. European Journal of Operational Research, 2007, 178 (3): 858 – 878.

[9] 刘勤明，李亚琴，吕文元，等．基于自适应隐式半马尔可夫模型的设备健康诊断与寿命预测方法 [J]．计算机集成制造系统，2016, 22 (09): 2187-2194. (Liu Qinming, Li Yaqin, Lyu Wenyuan, *et al.* Equipment health diagnostics and prognostics method based on adaptive HSMM [J]. Computer Integrated Manufacturing Systems, 2016, 22 (09): 2187-2194.)

[10] Chen A, Wu G S. Real time health prognosis and dynamic preventive maintenance policy for equipment under aging Markovian deterioration [J]. International Journal of Production Research. 2007, 45 (15): 3351-3379.

[11] 彭颖．基于退化隐式半马尔可夫模型的设备健康预测及系统性维护策略研究 [D]．上海：上海交通大学，2011. (Peng Ying. A Prognosis Method Using Age-Dependent Hidden Semi-Markov Model and A Systematic Maintenance Policy [D]. Shanghai: Shanghai Jiaotong University, 2011.)

[12] 何兆民，王少萍．基于时变状态转移隐半马尔可夫模型的寿命预测 [J]．湖南大学学报：自然科学版，2014, 41 (08): 47-53. (He Zhaoming, Wang Shaoping. Remaining Lifetime Prediction Based on Time-varying State Transition Probabilities of Hidden semi-Markov Model [j]. Journal of Hunan University: Natural Sciences, 2014, 41 (08): 47-53.)

[13] 原媛，卓东风．隐半马尔可夫模型在剩余寿命预测中的应用 [J]．计算机技术与发展，2014, 24 (01): 184-187+191. (Yuan Yuan, Zhuo Dongfeng. Application of Hidden Semi-Markov Model in Prediction of Residual Life [J]. Computer Technology and Development, 2014, 24 (01): 184-187+191.)

[14] 邬丽云，吴正鹏，齐英剑．基于遗传算法的灰色区间数的 GM (1, 1) 优化模型 [J]．控制与决策，2019, 34 (02): 445-448. (Wu Liyun, Wu Zhengpeng, Qi Yingjian. GM (1, 1) Model for Interval Grey Number Based on Genetic Algorithm [J]. Control and Decision, 2019, 34 (02): 445-448.)

[15] Krause P, Wasynczuk O, Sudhoff S, *et al.* Analysis of Electric Machinery and Drive Systems [M]. New York: John Wiley&Sons, lnc, 2013.

[16] 张小强，朱文辉，康铁宇，等．基于人工免疫算法的离散隐马尔可夫故障诊断模型优化 [J]．装备环境工程，2019, 16 (01): 63-67. (Zhang Xiaoqiang, Zhu Wenhui, Kang Tieyu, *et al.* Optimization of Discrete Hidden Markov Fault Diagnosis Model Based on Artificial Immune Algorithm [J]. Equipment Environmental Engineering, 2019, 16 (01): 63-67.)

[17] 张西宁，雷威，杨雨薇，等．采用自适应基因粒子群算法优化隐马尔可夫模型的方法及应用 [J]．西安交通大学学报，2018, 52 (08): 1-8. (Zhang Xining, Lei Wei, Yang Yuwei, *et al.* Adaptive Genetic Particle Swarm Algorithm for Optimization Hidden Markov Models with Applications [J]. Journal of Xi'an Jiaotong University, 2018, 52 (08): 1-8.)